

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Tormo, Mar
Tari, Ana M.
Lopez-Berestein, Gabriel

(ii) TITLE OF INVENTION: INHIBITION OF Bcl-2 PROTEIN EXPRESSION BY
LIPOSOMAL ANTISENSE OLIGODEOXYNUCLEOTIDES

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Arnold, White & Durkee
(B) STREET: P.O. Box 4433
(C) CITY: Houston
(D) STATE: Texas
(E) COUNTRY: United States of America
(F) ZIP: 77210

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US Unknown
(B) FILING DATE: Concurrently Herewith
(C) CLASSIFICATION: Unknown

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Wilson, Mark B.
(B) REGISTRATION NUMBER: 37,259
(C) REFERENCE/DOCKET NUMBER: UTXC:504

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (512) 418-3000
(B) TELEFAX: (512) 474-7577

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGCGTGCGC CATCCTTC

18

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACGGTCCGCC ACTCCTTCCC 20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGAAGGGCT TCTTCC 16

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5086 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1459..2175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGCCCGCCC CTCCGCGCCG CCTGCCCCGCC CGCCCGCCGC GCTCCCGCCC GCCGCTCTCC 60
GTGGCCCCGC CGCGCTGCCG CCGCCGCCGC TGCCAGCGAA GGTGCCGGGG CTCCGGGCCC 120
TCCCTGCCGG CGGCCGTCAG CGCTCGGAGC GAACTGCGCG ACGGGAGGTC CGGGAGGCGA 180
CCGTAGTCGC GCCGCCGCGC AGGACCAGGA GGAGGAGAAA GGGTGCGCAG CCCGGAGGCG 240
GGGTGCGCCG GTGGGGTGCA GCGGAAGAGG GGGTCCAGGG GGGAGAACTT CGTAGCAGTC 300
ATCCTTTTTA GGAAAAGAGG GAAAAAATAA AACCTCCCC CACCACCTCC TTCTCCCCAC 360

	CCCTCGCCGC ACCACACACA GCGCGGGCTT CTAGCGCTCG GCACCGGCGG GCCAGGCGCG	420
	TCCTGCCTTC ATTTATCCAG CAGCTTTTCG GAAAATGCAT TTGCTGTTTCG GAGTTTAATC	480
5	AGAAGACGAT TCCTGCCTCC GTCCCCGGCT CCTTCATCGT CCCATCTCCC CTGTCTCTCT	540
	CCTGGGGAGG CGTGAAGCGG TCCCGTGGAT AGAGATTCAT GCCTGTGTCC GCGCGTGTGT	600
10	GCGCGCGTAT AAATTGCCGA GAAGGGGAAA ACATCACAGG ACTTCTGCGA ATACCGGACT	660
	GAAAATTGTA ATTCATCTGC CGCCGCCGCT GCCAAAAAAA AACTCGAGCT CTTGAGATCT	720
	CCGGTTGGGA TTCCTGCGGA TTGACATTTT TGTGAAGCAG AAGTCTGGGA ATCGATCTGG	780
15	AAATCCTCCT AATTTTTACT CCCTCTCCCC CCGACTCCTG ATTCATTGGG AAGTTTCAAA	840
	TCAGCTATAA CTGGAGAGTG CTGAAGATTG ATGGGATCGT TGCCTTATGC ATTTGTTTTG	900
20	GTTTTACAAA AAGGAAACTT GACAGAGGAT CATGCTGTAC TTAAAAAATA CAAGTAAGTC	960
	TCGCACAGGA AATTGGTTTA ATGTAACTTT CAATGGAAAC CTTTGAGATT TTTTACTTAA	1020
	AGTGCATTCG AGTAAATTTA ATTTCCAGGC AGCTTAATAC ATTGTTTTTA GCCGTGTTAC	1080
25	TTGTAGTGTG TATGCCCTGC TTTCACTCAG TGTGTACAGG GAAACGCACC TGATTTTTTA	1140
	CTTATTAGTT TGTTTTTTCT TTAACCTTTC AGCATCACAG AGGAAGTAGA CTGATATTAA	1200
30	CAATACTTAC TAATAATAAC GTGCCTCATG AAATAAAGAT CCGAAAGGAA TTGGAATAAA	1260
	AATTCCTGC GTCTCATGCC AAGAGGGAAA CACCAGAATC AAGTGTTCCG CGTGATTGAA	1320
	GACACCCCT CGTCCAAGAA TGCAAAGCAC ATCCAATAAA ATAGCTGGAT TATAACTCCT	1380
35	CTTCTTTCTC TGGGGGCCGT GGGGTGGGAG CTGGGGCGAG AGGTGCCGTT GGCCCCCGTT	1440
	GCTTTTCCTC TGGGAAGG ATG GCG CAC GCT GGG AGA ACG GGG TAC GAC AAC	1491
	Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn	
	1 5 10	
40	CGG GAG ATA GTG ATG AAG TAC ATC CAT TAT AAG CTG TCG CAG AGG GGC	1539
	Arg Glu Ile Val Met Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly	
	15 20 25	
45	TAC GAG TGG GAT GCG GGA GAT GTG GGC GCC GCG CCC CCG GGG GCC GCC	1587
	Tyr Glu Trp Asp Ala Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala	
	30 35 40	
50	CCC GCA CCG GGC ATC TTC TCC TCC CAG CCC GGG CAC ACG CCC CAT CCA	1635
	Pro Ala Pro Gly Ile Phe Ser Ser Gln Pro Gly His Thr Pro His Pro	
	45 50 55	

	GCC GCA TCC CGC GAC CCG GTC GCC AGG ACC TCG CCG CTG CAG ACC CCG	1683
	Ala Ala Ser Arg Asp Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro	
	60 65 70 75	
5	GCT GCC CCC GGC GCC GCC GCG GGG CCT GCG CTC AGC CCG GTG CCA CCT	1731
	Ala Ala Pro Gly Ala Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro	
	80 85 90	
10	GTG GTC CAC CTG GCC CTC CGC CAA GCC GGC GAC GAC TTC TCC CGC CGC	1779
	Val Val His Leu Ala Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg	
	95 100 105	
15	TAC CGC GGC GAC TTC GCC GAG ATG TCC AGC CAG CTG CAC CTG ACG CCC	1827
	Tyr Arg Gly Asp Phe Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro	
	110 115 120	
20	TTC ACC GCG CGG GGA CGC TTT GCC ACG GTG GTG GAG GAG CTC TTC AGG	1875
	Phe Thr Ala Arg Gly Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg	
	125 130 135	
25	GAC GGG GTG AAC TGG GGG AGG ATT GTG GCC TTC TTT GAG TTC GGT GGG	1923
	Asp Gly Val Asn Trp Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly	
	140 145 150 155	
30	GTC ATG TGT GTG GAG AGC GTC AAC CGG GAG ATG TCG CCC CTG GTG GAC	1971
	Val Met Cys Val Glu Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp	
	160 165 170	
35	AAC ATC GCC CTG TGG ATG ACT GAG TAC CTG AAC CGG CAC CTG CAC ACC	2019
	Asn Ile Ala Leu Trp Met Thr Glu Tyr Leu Asn Arg His Leu His Thr	
	175 180 185	
40	TGG ATC CAG GAT AAC GGA GGC TGG GAT GCC TTT GTG GAA CTG TAC GGC	2067
	Trp Ile Gln Asp Asn Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly	
	190 195 200	
45	CCC AGC ATG CGG CCT CTG TTT GAT TTC TCC TGG CTG TCT CTG AAG ACT	2115
	Pro Ser Met Arg Pro Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr	
	205 210 215	
50	CTG CTC AGT TTG GCC CTG GTG GGA GCT TGC ATC ACC CTG GGT GCC TAT	2163
	Leu Leu Ser Leu Ala Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr	
	220 225 230 235	
55	CTG AGC CAC AAG TGAAGTCAAC ATGCCTGCCC CAAACAAATA TGCAAAAGGT	2215
	Leu Ser His Lys	
60	TCACTAAAGC AGTAGAAATA ATATGCATTG TCAGTGATGT ACCATGAAAC AAAGCTGCAG	2275
	GCTGTTTAAG AAAAAATAAC ACACATATAA ACATCACACA CACAGACAGA CACACACACA	2335
	CACAACAATT AACAGTCTTC AGGCAAAACG TCGAATCAGC TATTTACTGC CAAAGGGAAA	2395

	TATCATTAT TTTTACATT ATTAAGAAAA AAGATTTATT TATTTAAGAC AGTCCCATCA	2455
	AAACTCCGTC TTTGGAAATC CGACCACTAA TTGCCAAACA CCGCTTCGTG TGGCTCCACC	2515
5	TGGATGTTCT GTGCCTGTAA ACATAGATTC GCTTTCCATG TTGTTGGCCG GATCACCATC	2575
	TGAAGAGCAG ACGGATGGAA AAAGGACCTG ATCATTGGGG AAGCTGGCTT TCTGGCTGCT	2635
10	GGAGGCTGGG GAGAAGGTGT TCATTCACTT GCATTTCTTT GCCCTGGGGG CGTGATATTA	2695
	ACAGAGGGAG GGTTCCTCGT GGGGGAAGTC CATGCCTCCC TGGCCTGAAG AAGAGACTCT	2755
	TTGCATATGA CTCACATGAT GCATACCTGG TGGGAGGAAA AGAGTTGGGA ACTTCAGATG	2815
15	GACCTAGTAC CCACTGAGAT TTCCACGCCG AAGGACAGCG ATGGGAAAAA TGCCCTTAAA	2875
	TCATAGGAAA GTATTTTTTT AAGCTACCAA TTGTGCCGAG AAAAGCATTT TAGCAATTTA	2935
20	TACAATATCA TCCAGTACCT TAAACCCTGA TTGTGTATAT TCATATATTT TGGATACGCA	2995
	CCCCCAACT CCCAATACTG GCTCTGTCTG AGTAAGAAAC AGAATCCTCT GGAACCTGAG	3055
	GAAGTGAACA TTTCGGTGAC TTCCGATCAG GAAGGCTAGA GTTACCCAGA GCATCAGGCC	3115
25	GCCACAAGTG CCTGCTTTTA GGAGACCGAA GTCCGCAGAA CCTACCTGTG TCCAGCTTG	3175
	GAGGCCTGGT CCTGGAACTG AGCCGGGCCC TCACTGGCCT CCTCCAGGGA TGATCAACAG	3235
30	GGTAGTGTGG TCTCCGAATG TCTGGAAGCT GATGGATGGA GCTCAGAATT CCACTGTCAA	3295
	GAAAGAGCAG TAGAGGGGTG TGGCTGGGCC TGTACCCTG GGGCCCTCCA GGTAGCCCCG	3355
	TTTTCACGTG GAGCATAGGA GCCACGACCC TTCTTAAGAC ATGTATCACT GTAGAGGGAA	3415
35	GGAACAGAGG CCCTGGGCCT TCCTATCAGA AGGACATGGT GAAGGCTGGG AACGTGAGGA	3475
	GAGGCAATGG CCACGGCCCA TTTTGGCTGT AGCACATGGC ACGTTGGCTG TGTGGCCTTG	3535
40	GCCACCTGTG AGTTTAAAGC AAGGCTTTAA ATGACTTTGG AGAGGGTCAC AAATCCTAAA	3595
	AGAAGCATTG AAGTGAGGTG TCATGGATTA ATTGACCCT GTCTATGGAA TTACATGTAA	3655
	AACATTATCT TGTCACGTGA GTTTGGTTTT ATTTGAAAAC CTGACAAAAA AAAAGTTCCA	3715
45	GGTGTGGAAT ATGGGGGTTA TCTGTACATC CTGGGGCATT AAAAAAAT CAATGGTGGG	3775
	GAACATAAAA GAAGTAACAA AAGAAGTGAC ATCTTCAGCA AATAAACTAG GAAATTTTTT	3835
50	TTTCTTCCAG TTTAGAATCA GCCTTGAAAC ATTGATGGAA TAACTCTGTG GCATTATTGC	3895
	ATTATATACC ATTTATCTGT ATTAACCTTG GAATGTACTC TGTTCAATGT TTAATGCTGT	3955
	GGTTGATATT TCGAAAGCTG CTTTAAAAAA ATACATGCAT CTCAGCGTTT TTTTGTTTTT	4015

2672631.013197

AATTGTATTT AGTTATGGCC TATACACTAT TTGTGAGCAA AGGTGATCGT TTTCTGTTTG 4075
AGATTTTAT CTCTTGATTC TTCAAAAGCA TTCTGAGAAG GTGAGATAAG CCCTGAGTCT 4135
5 CAGCTACCTA AGAAAAACCT GGATGTCACT GGCCACTGAG GAGCTTTGTT TCAACCAAGT 4195
CATGTGCATT TCCACGTCAA CAGAATTGTT TATTGTGACA GTTATATCTG TTGTCCCTTT 4255
10 GACCTTGTTT CTTGAAGGTT TCCTCGTCCC TGGGCAATTC CGCATTTAAT TCATGGTATT 4315
CAGGATTACA TGCATGTTTG GTTAAACCCA TGAGATTCAT TCAGTTAAAA ATCCAGATGG 4375
CGAATGACCA GCAGATTCAA ATCTATGGTG GTTTGACCTT TAGAGAGTTG CTTTACGTGG 4435
15 CCTGTTTCAA CACAGACCCA CCCAGAGCCC TCCTGCCCTC CTTCCGCGGG GGCTTTCTCA 4495
TGGCTGTCCT TCAGGGTCTT CCTGAAATGC AGTGGTCGTT ACGCTCCACC AAGAAAGCAG 4555
GAAACCTGTG GTATGAAGCC AGACCTCCCC GCGGGGCTC AGGGAACAGA ATGATCAGAC 4615
20 CTTTGAATGA TTCTAATTTT TAAGCAAAAT ATTATTTTAT GAAAGGTTTA CATTGTCAAA 4675
GTGATGAATA TGGAATATCC AATCCTGTGC TGCTATCCTG CCAAATCAT TTTAATGGAG 4735
25 TCAGTTTGCA GTATGCTCCA CGTGGTAAGA TCCTCCAAGC TGCTTTAGAA GTAACAATGA 4795
AGAACGTGGA CGTTTTTAAT ATAAAGCCTG TTTTGTCTTT TGTTGTTGTT CAAACGGGAT 4855
TCACAGAGTA TTTGAAAAAT GTATATATAT TAAGAGGTCA CGGGGGCTAA TTGCTAGCTG 4915
30 GCTGCCTTTT GCTGTGGGGT TTTGTTACCT GGTTTTAATA ACAGTAAATG TGCCAGCCT 4975
CTTGCCCCA GAACTGTACA GTATTGTGGC TGCACTTGCT CTAAGAGTAG TTGATGTTGC 5035
35 ATTTTCCTTA TTGTTAAAAA CATGTTAGAA GCAATGAATG TATATAAAAG C 5086

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 239 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
1 5 10 15
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
20 25 30

Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
 35 40 45
 Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
 5 50 55 60
 Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
 65 70 75 80
 Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
 10 85 90 95
 Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
 15 100 105 110
 Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
 115 120 125
 Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
 20 130 135 140
 Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
 145 150 155 160
 Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
 25 165 170 175
 Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
 30 180 185 190
 Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro
 195 200 205
 Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala
 35 210 215 220
 Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu Ser His Lys
 225 230 235

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 911 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 147..761

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

082511.01319

	TGATTGAAGA CACCCCCTCG TCCAAGAATG CAAAGCACAT CCAATAAAAT AGCTGGATTA	60
	TAACCTCTCT TCTTTCTCTG GGGGCCGTGG GGTGGGAGCT GGGGCGAGAG GTGCCGTTGG	120
5	CCCCCGTTGC TTTTCCTCTG GGAAGG ATG GCG CAC GCT GGG AGA ACG GGG TAC Met Ala His Ala Gly Arg Thr Gly Tyr 1 5	173
10	GAC AAC CGG GAG ATA GTG ATG AAG TAC ATC CAT TAT AAG CTG TCG CAG Asp Asn Arg Glu Ile Val Met Lys Tyr Ile His Tyr Lys Leu Ser Gln 10 15 20 25	221
15	AGG GGC TAC GAG TGG GAT GCG GGA GAT GTG GGC GCC GCG CCC CCG GGG Arg Gly Tyr Glu Trp Asp Ala Gly Asp Val Gly Ala Ala Pro Pro Gly 30 35 40	269
20	GCC GCC CCC GCA CCG GGC ATC TTC TCC TCC CAG CCC GGG CAC ACG CCC Ala Ala Pro Ala Pro Gly Ile Phe Ser Ser Gln Pro Gly His Thr Pro 45 50 55	317
25	CAT CCA GCC GCA TCC CGC GAC CCG GTC GCC AGG ACC TCG CCG CTG CAG His Pro Ala Ala Ser Arg Asp Pro Val Ala Arg Thr Ser Pro Leu Gln 60 65 70	365
30	ACC CCG GCT GCC CCC GGC GCC GCC GCG GGG CCT GCG CTC AGC CCG GTG Thr Pro Ala Ala Pro Gly Ala Ala Ala Gly Pro Ala Leu Ser Pro Val 75 80 85	413
35	CCA CCT GTG GTC CAC CTG GCC CTC CGC CAA GCC GGC GAC GAC TTC TCC Pro Pro Val Val His Leu Ala Leu Arg Gln Ala Gly Asp Asp Phe Ser 90 95 100 105	461
40	CGC CGC TAC CGC GGC GAC TTC GCC GAG ATG TCC AGC CAG CTG CAC CTG Arg Arg Tyr Arg Gly Asp Phe Ala Glu Met Ser Ser Gln Leu His Leu 110 115 120	509
45	ACG CCC TTC ACC GCG CGG GGA CGC TTT GCC ACG GTG GTG GAG GAG CTC Thr Pro Phe Thr Ala Arg Gly Arg Phe Ala Thr Val Val Glu Glu Leu 125 130 135	557
50	TTC AGG GAC GGG GTG AAC TGG GGG AGG ATT GTG GCC TTC TTT GAG TTC Phe Arg Asp Gly Val Asn Trp Gly Arg Ile Val Ala Phe Phe Glu Phe 140 145 150	605
	GGT GGG GTC ATG TGT GTG GAG AGC GTC AAC CGG GAG ATG TCG CCC CTG Gly Gly Val Met Cys Val Glu Ser Val Asn Arg Glu Met Ser Pro Leu 155 160 165	653
	GTG GAC AAC ATC GCC CTG TGG ATG ACT GAG TAC CTG AAC CGG CAC CTG Val Asp Asn Ile Ala Leu Trp Met Thr Glu Tyr Leu Asn Arg His Leu 170 175 180 185	701

CAC ACC TGG ATC CAG GAT AAC GGA GGC TGG GTA GGT GCA TCT GGT GAT 749
 His Thr Trp Ile Gln Asp Asn Gly Gly Trp Val Gly Ala Ser Gly Asp
 190 195 200

5 GTG AGT CTG GGC TGAGGCCACA GGTCCGAGAT CGGGGGTTGG AGTGCGGGTG 801
 Val Ser Leu Gly
 205

10 GGCTCCTGGG CAATGGGAGG CTGTGGAGCC GGCGAAATAA AATCAGAGTT GTTGCTTCCC 861
 GGCGTGTCCC TACCTCCTCC TCTGGACAAA GCGTTCCTC CCAACCTGAC 911

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 205 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
 1 5 10 15
 Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
 20 25 30
 Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
 35 40 45
 Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
 50 55 60
 Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
 65 70 75 80
 Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
 85 90 95
 Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
 100 105 110
 Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
 115 120 125
 Arg Phe Ala Thr Val Val Glu Leu Phe Arg Asp Gly Val Asn Trp
 130 135 140
 Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
 145 150 155 160

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
165 170 175

5 Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
180 185 190

Gly Gly Trp Val Gly Ala Ser Gly Asp Val Ser Leu Gly
195 200 205

08726211.013197